

# Scientific Data Formats for Space Weather Model Data Management



August 1, 2001

## *A Case for HDF5*

Paula J. Reitan

NASA/GSFC

CCMC

[Paula.J.Reitan.1@gsfc.nasa.gov](mailto:Paula.J.Reitan.1@gsfc.nasa.gov)

<http://ccmc.gsfc.nasa.gov/~preitan>



Recommended Reference:

NASA/GSFC 586/587 Science Data Processing Portal

<http://that.gsfc.nasa.gov/gss/>

## Advantages of Scientific Data Formats

- Output from space weather models is typically stored on disk for further analysis and visualization. Why not use a standard scientific data format?
  - Utilities available.
  - Interact with various commercial and open-source software.
  - Compression available.
  - Self-describing using metadata.
  - Discipline/problem/application/platform independent.
  - Access data/metadata using high-level APIs.
  - Data between different research groups and applications can be shared in a format independent fashion.



## Scientific Data Formats Considered

- **Common Data Format (CDF)**
  - Sponsored by National Space Science Data Center (NSSDC) at NASA/GSFC.
  - <http://nssdc.gsfc.nasa.gov/cdf/>
- **Hierarchical Data Format (HDF)**
  - Sponsored by National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign.
  - <http://hdf.ncsa.uiuc.edu/>
  - NCSA recommends using HDF5, especially if you are a new user and are not constrained to using HDF4.

3

August 1, 2001



## CDF vs HDF

- Both self-describing using metadata.
- Both discipline/problem/application/platform independent.
- Both access data/metadata using high-level APIs.
- Both binary formats and efficiently store data. Both support compression.
- Both open-source.
- CDF has more utilities/tools.
- CDF is supported by more commercial and open-source software.
- HDF5 is better documented.
- HDF5 has many more predefined data types.
- HDF5 is a newer, simpler, more general data format.
- HDF5 supports storage of hierarchical data via **groups**.

4

August 1, 2001

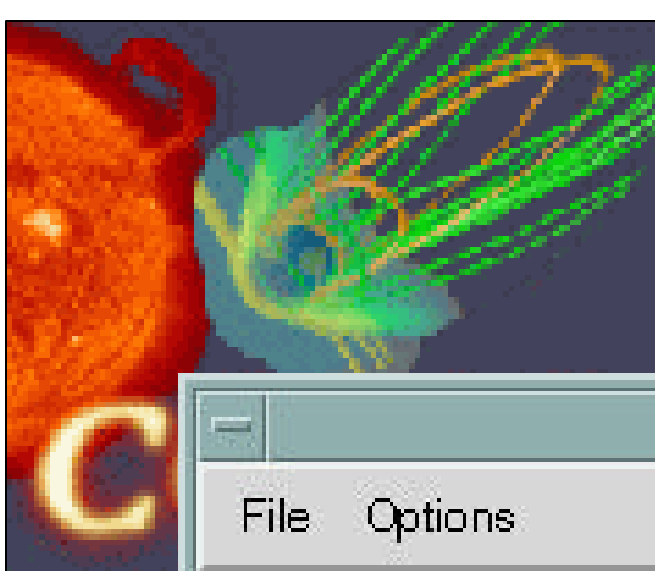


## Space Weather View (SWV)

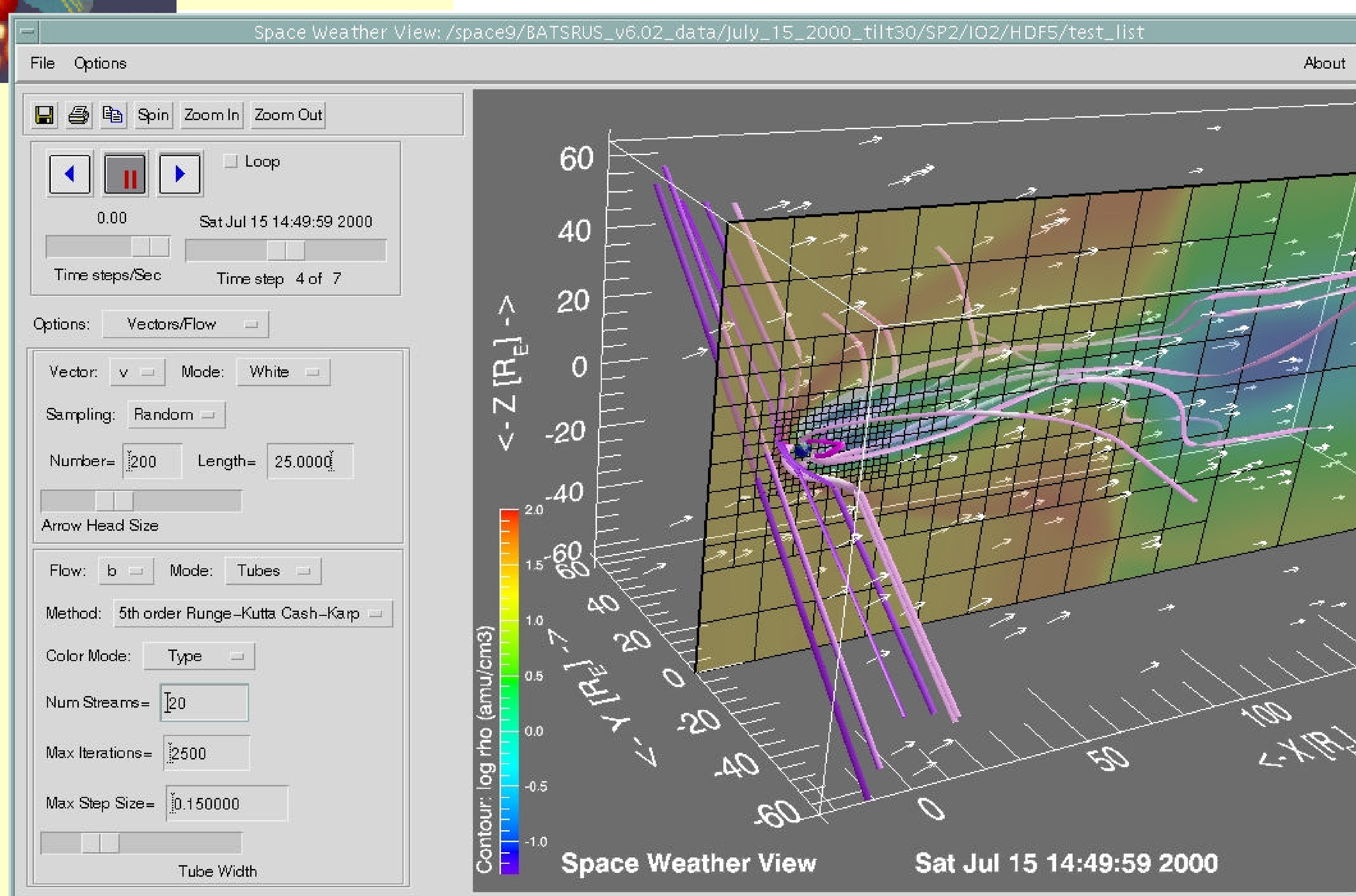
- Space Weather View (SWV) is a program written using **IDL** object graphics to visualize in 3D output from space weather models.
- SWV currently displays output from **BATS-R-US** which has been converted to **HDF5**.
  - Structure of this HDF5 file is specific to BATS-R-US and should be made more general.
- Plan to incorporate display of output from other models studied by the CCMC into SWV.

5

August 1, 2001



## SWV Screen Shot



6

August 1, 2001



## Benefits of Using HDF5 for SWV

- Hierarchical structure of Gombosi's BATS-R-US MHD adaptive mesh (octree) can be easily stored using the **group** feature of HDF5.
  - See <http://ccmc.gsfc.nasa.gov/~preitan/CCMC/SWV>
- C, C++, F90 and Java HDF5 API available. Very well documented and easy to learn.
- Size of HDF5 file is slightly smaller than the current (.idl and .out) adhoc binary file formats. Compression has not been tested yet.
- The **h5dump** and **h5ls** utilities can be used to obtain information about the HDF5 file.
- Parallel HDF5 library available for parallel access to HDF5 files (has not been tested yet).

7

August 1, 2001



## Disadvantages of Using HDF5 for SWV

- IDL does not currently support HDF5.
  - Work-around: make external function calls to shared library.
    - Memory is allocated in IDL and a pointer passed to an external C++ function. C++ function makes HDF5 API calls to read requested data into memory allocated by IDL.

8

August 1, 2001



## Recommendations

- Establish a set of guidelines for describing and storing space weather model data using a scientific data format.
  - Guidelines established for ISTP/IACG CDF files would be a good starting point ([http://nssdc.gsfc.nasa.gov/space/spdf/istp\\_guide/istp\\_guide.html](http://nssdc.gsfc.nasa.gov/space/spdf/istp_guide/istp_guide.html)).
- The goal is to make the resulting data file correctly and independently usable by the space weather modeling community.
- Enable analysis and visualization tools to be written once and used independent of researcher, application, or platform.